

CLAIMS.

We claim

1. A method of pre-processing of a machine-readable form image with non-fixed fields layout, comprising

- a bit-mapped image of the filled in form,
- at least one model of a form, containing spatial and parametric properties of objects thereof,
- preliminarily assigning at least one object of the form as a reference point for spatial binding of at least one data input field thereof,
- performing at least the following steps:
 - a step of eliminating skew, distortion and noise of the image,
 - a step of parsing the image into regions,
 - a step of defining location of at least one data input field relatively to at least one reference point, said step further comprising:
 - selecting a data input field to search in the said form model description,
 - accepting from the model description at least one reference point properties for the spatial relative reference of the said data input field,
 - searching said at least one reference point on the form image,
 - searching the said data input field on the form image relatively to at least one reference point taking into account all spatial and parametrical properties, described in the form model,
 - profound identification of the said data input field position in the case of multiple search result.

2. The method as recited in claim 1, wherein the said reference point is represented by a text region.

3. The method as recited in claim 2, wherein the said text region is additionally subjected to a text recognition.
4. The method as recited in claim 1, wherein in the case of multiple search result the said identification is performed via setting up and accepting of hypotheses and compliance estimation with form model description.
5. The method as recited in claim 4, wherein an additional information about the said data input field is used.
6. The method as recited in claim 1, wherein the data input field may be used as a reference point.
7. The method as recited in claim 1, wherein the said step of data input field identification in a case of multiple identification result is performed partly manually.
8. The method as recited in claim 1, wherein the reference point spatial location is not fixed.
9. The method as recited in claim 1, wherein one reference point is used for spatial binding of more then one data input field.
10. The method as recited in claim 1, wherein the spatial binding of one data input field is performed to more then one reference point.
11. The method as recited in claim 1, wherein a reference point comprises more then one form object.
12. The method as recited in claim 1, wherein a reference point is described in a form of alternative.